

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: August 27, 2004, 21:39:31 ; Search time 21 Seconds  
(without alignments)  
501.509 Million cell updates/sec

Title: US-10-633-202-2\_COPY\_399\_602

Perfect score: 1192

Sequence: 1 PYPGNIPLKCYCGNLVVEEG.....QFHLNDTTCWGTDYHLGNWAI 204

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 389414 seqs, 51625971 residues

Total number of hits satisfying chosen parameters: 389414

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents AA.\*

- 1: /cgn2\_6/ptodata/2/iaa/5A COMB.pap.\*
- 2: /cgn2\_6/ptodata/2/iaa/5B COMB.pap.\*
- 3: /cgn2\_6/ptodata/2/iaa/6A COMB.pap.\*
- 4: /cgn2\_6/ptodata/2/iaa/6B COMB.pap.\*
- 5: /cgn2\_6/ptodata/2/iaa/PCTUS COMB.pap.\*
- 6: /cgn2\_6/ptodata/2/iaa/backfiles1.pap.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	864	72.5	722	4	US-09-617-145-2
2	431.5	36.2	751	2	US-08-836-443-3
3	422.5	35.4	787	4	US-09-548-797B-5
4	422.5	35.4	802	4	US-09-632-098-2
5	422.5	35.4	812	4	US-09-632-098-4
6	422.5	35.4	849	4	US-09-548-797B-6
7	401.5	33.7	814	3	US-09-813-819-4
8	401.5	33.7	814	4	US-09-920-048-4
9	401.5	33.7	855	3	US-09-813-819-2
10	401.5	33.7	855	4	US-09-920-048-2
11	390	32.7	651	1	US-08-264-101-2
12	390	32.7	651	2	US-08-765-243-2
13	390	32.7	651	5	PCT-US95-07295-2
14	390	32.7	734	2	US-08-765-243-8
15	390	32.7	734	5	PCT-US95-07295-8
16	384.5	32.3	696	3	US-09-351-414-2
17	384	32.2	670	1	US-08-243-542-3
18	384	32.2	670	1	US-08-477-407-3
19	384	32.2	670	1	US-08-484-355-3
20	384	32.2	769	1	US-08-243-542-4
21	384	32.2	769	1	US-08-477-407-4
22	384	32.2	769	1	US-08-484-355-4
23	368	30.9	775	4	US-09-786-256C-15
24	368	30.9	775	4	US-09-786-256C-32
25	359	30.1	457	1	US-08-264-101-4
26	359	30.1	457	2	US-08-765-243-4
27	359	30.1	457	5	PCT-US95-07295-4

28	359	30.1	735	2	US-08-765-243-6
29	359	30.1	735	5	PCT-US95-07295-6
30	347	29.1	524	1	US-08-243-542-2
31	347	29.1	524	1	US-08-477-407-2
32	347	29.1	524	1	US-08-484-355-2
33	345	28.9	746	4	US-09-548-797B-4
34	344	28.9	488	1	US-08-243-542-1
35	344	28.9	488	1	US-08-477-407-1
36	344	28.9	488	1	US-08-484-355-1
37	342.5	28.7	781	4	US-09-738-946-8
38	332.5	27.9	616	4	US-09-608-790-1
39	312	26.2	613	4	US-09-026-001A-10
40	312	26.2	621	4	US-09-026-001A-18
41	306.5	25.7	621	4	US-09-026-001A-6
42	283.5	23.8	540	4	US-09-786-256C-1
43	283.5	23.8	540	4	US-09-786-256C-30
44	272	22.8	529	2	US-08-836-442-3
45	270.5	22.7	462	4	US-09-026-001A-16

ALIGNMENTS

RESULT 1

US-09-617-145-2

; Sequence 2, Application US/09617145

; Patent No. 6485956

; GENERAL INFORMATION:

; APPLICANT: Cerretti, Douglas P.

; TITLE OF INVENTION: SVPH1-8 DNA and Polypeptides

; FILE REFERENCE: 03260.0050-00304

; CURRENT APPLICATION NUMBER: US/09/617,145

; CURRENT FILING DATE: 2000-07-14

; PRIOR APPLICATION NUMBER: 60/071,505

; PRIOR FILING DATE: 1998-01-14

; NUMBER OF SEQ ID NOS: 2

; SOFTWARE: Patentin Ver. 2.0

; SEQ ID NO 2

; LENGTH: 722

; TYPE: PRT

; ORGANISM: Homo sapiens

US-09-617-145-2

Query Match 72.5%; Score 864; DB 4; Length 722;  
Best Local Similarity 70.6%; Pred. No. 1.2e-63;  
Matches 144; Conservative 24; Mismatches 36; Indels 0; Gaps 0;

Qy	1	PYPGNIPLKCYCGNLVVEEGECDCGTIRQCAKDPCCLLNCTLHPGACAFGLCKCKCF	60
Db	398	PRLGEIFMLKRCGNGVVEREBQDCGSCVQCEQDACLLNCTLRFGAACAFGLCKCKCF	457
Qy	61	LPSGTLCRQVQGECDLPWCNGTSHQCDDVYVQDGISCNVNAFCYEKTCNNHDIQCKEI	120
Db	458	MPSGELCQFVNECDLPWCNGTSHQCDDVYVQDGIPCSDSAYCYKRCNNHDIQCKEI	517
Qy	121	FGQDARSASQCYQBINTQGRFGRHCGIVGTYVVKWTPDIMGRCVQCNVGVIPNLIEH	180
Db	518	FGDAKASQNCYKEINSQGRFGRHCGIVGTYVVKWTPDIMGRCVQCNVGVIPNLIEH	577
Qy	181	STVQGFHLNDTTCWGTDYHLGNWAI	204
Db	578	FTLQTHINGVTCMGIDYHLRMI	601

RESULT 2

US-08-836-443-3

; Sequence 3, Application US/08836443

; Patent No. 5883241

; GENERAL INFORMATION:

; APPLICANT: DOCHERTY, Andrew, J.P

; APPLICANT: SLOCOMBE, Patrick, M.

; TITLE OF INVENTION: DNA SEQUENCES CODING FOR A

; TITLE OF INVENTION: HUMAN METALLOPROTEINASE AND VARIANTS THEREOF



Result No.	Score	Query			DB	ID	Description
		Match	Length	%			
1	1192	100.0	726	2	AAV17413	Human SVP	
2	1192	100.0	726	2	AAV03223	Amino aci	
3	1192	100.0	726	7	ADC78877	Human PRO	
4	1192	100.0	776	6	ABR39425	Human GEN	
5	1128	94.6	535	2	AAV17414	SVP#1-26	
6	1128	94.6	535	4	AAE13056	ADAM-20d1	
7	864	72.5	542	2	AAV03224	Amino aci	
8	864	72.5	722	2	AAV28635	Human SVP	
9	864	70.5	722	6	ABG76200	Human sna	
10	840	70.5	523	4	AAE13057	ADAM-21d1	
11	788	66.1	766	3	AAV07739	A snake v	
12	788	66.1	787	3	AAV07730	A snake v	
13	788	66.1	820	3	AAV07741	A snake v	
14	773	64.8	528	4	AAE13061	ADAM-29d1	
15	590.5	49.5	499	3	AAV07738	A snake v	
16	590.5	49.5	781	3	AAV07743	A snake v	
17	590.5	49.5	790	3	AAV07742	A snake v	
18	590.5	49.5	790	3	AAV07705	Amino aci	
19	590.5	49.5	790	4	AAU12273	Human PRO	
20	590.5	49.5	790	6	ABO17717	Novel hum	
21	590.5	49.5	790	6	ABU80971	Human PRO	
22	590.5	49.5	790	6	ABU66671	Human PRO	
23	590.5	49.5	790	6	ABU59752	Novel sec	
24	590.5	49.5	790	6	ABO24942	Human sec	
25	590.5	49.5	790	6	ABU66947	Human sec	

CC	prognosis of testicular cancers	
XX		
SQ	Sequence 726 AA;	SQ Sequence 726 AA;
	Query Match 100.0%; Score 1192; DB 2; Length 726;	Query Match 100.0%; Score 1192; DB 2; Length 726;
	Best Local Similarity 100.0%; Pred. No. 8.4e-88;	Best Local Similarity 100.0%; Pred. No. 8.4e-88;
	Matches 204; Conservative 0; Mismatches 0; Indels 0; Gaps 0;	Matches 204; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY	1 PYPGNIFRLKYCGNLVVEEGECDCGTIRQCAKDPCCLLNCTLHPGAACAFGICCKDKCF 60	1 PYPGNIFRLKYCGNLVVEEGECDCGTIRQCAKDPCCLLNCTLHPGAACAFGICCKDKCF 60
Db	399 PYPGNIFRLKYCGNLVVEEGECDCGTIRQCAKDPCCLLNCTLHPGAACAFGICCKDKCF 458	399 PYPGNIFRLKYCGNLVVEEGECDCGTIRQCAKDPCCLLNCTLHPGAACAFGICCKDKCF 458
QY	61 LPSGTLCRQVGECDLPWCNCTSHQCPDDVVQDGISCNVNAFCYEKTCNNHDIQCKEI 120	61 LPSGTLCRQVGECDLPWCNCTSHQCPDDVVQDGISCNVNAFCYEKTCNNHDIQCKEI 120
Db	459 LPSGTLCRQVGECDLPWCNCTSHQCPDDVVQDGISCNVNAFCYEKTCNNHDIQCKEI 518	459 LPSGTLCRQVGECDLPWCNCTSHQCPDDVVQDGISCNVNAFCYEKTCNNHDIQCKEI 518
QY	121 FQDARSASQSCYQIEINTQGNRFGHCGIVGTTTYYKWTPTDIMCGRVQCNVGVIPNLIEH 180	121 FQDARSASQSCYQIEINTQGNRFGHCGIVGTTTYYKWTPTDIMCGRVQCNVGVIPNLIEH 180
Db	519 FQDARSASQSCYQIEINTQGNRFGHCGIVGTTTYYKWTPTDIMCGRVQCNVGVIPNLIEH 578	519 FQDARSASQSCYQIEINTQGNRFGHCGIVGTTTYYKWTPTDIMCGRVQCNVGVIPNLIEH 578
QY	181 STVQOFHLNDTTCWGTDYHLGMAI 204	181 STVQOFHLNDTTCWGTDYHLGMAI 204
Db	579 STVQOFHLNDTTCWGTDYHLGMAI 602	579 STVQOFHLNDTTCWGTDYHLGMAI 602
	RESULT 2	RESULT 3
ID	AA03223	ADC7877
XX	AA03223 standard; protein; 726 AA.	ID ADC7877 standard; protein; 726 AA.
AC	AA03223;	XX AC ADC7877;
XX		XX AC ADC7877;
DT	16-AUG-1999 (first entry)	DT 01-JAN-2004 (first entry)
XX		XX Human PRO protein #53.
DE	Amino acid sequence of the novel metalloprotease ADAM 16a.	XX human; PRO; cancer; inflammatory bowel disease; ulcerative colitis;
XX	membrane metalloprotease protein; human; ADAM 16a; testis; contraceptive;	KW Crohn's disease.
KW	vaccine; fertility diagnosis; fertility treatment.	XX Homo sapiens.
XX		OS Homo sapiens.
OS	Homo sapiens.	XX WO2003034984-A2.
XX		XX 01-MAY-2003.
FH	Key Location/Qualifiers	PD 15-OCT-2002; 2002WO-US033070.
FT	Peptide 1..31	PF 19-OCT-2001; 2001US-0340083P.
FT	/note= "signal peptide"	PR (GETH ) GENENTECH INC.
FT	32..727	XX PA Goddard A, Gurney AL;
FT	/note= "mature protein"	XX PI WPI; 2003-481990/45.
XX		XX DR N-PSDB; ADC78876.
PN	WO9907856-A1.	XX PT New PRO polynucleotide and polypeptide, useful for the manufacture of a
XX		PT medicament for diagnosing or treating cancer or inflammatory bowel
PD	18-FEB-1999.	PT disorder e.g., ulcerative colitis or Crohn's disease.
XX		PS Claim 12; SEQ ID NO 106; 327pp; English.
XX		XX The invention comprises the amino acid and coding sequences of human PRO
PF	05-AUG-1998; 98WO-EP004859.	CC proteins. The DNA and protein sequences of the invention are useful for
XX		CC the diagnosis and treatment of cancer and inflammatory bowel disease
XX		CC (e.g. ulcerative colitis or Crohn's disease). The present amino acid
XX		CC sequence represents a human PRO protein of the invention.
PA	(GLAX ) GLAXO GROUP LTD.	XX SQ Sequence 726 AA;
XX		XX Query Match 100.0%; Score 1192; DB 7; Length 726;
PI	Hoof Van Huijsduinen RAM;	XX Best Local Similarity 100.0%; Pred. No. 8.4e-88;
XX		XX Matches 204; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX		XX 1 PYPGNIFRLKYCGNLVVEEGECDCGTIRQCAKDPCCLLNCTLHPGAACAFGICCKDKCF 60
DR	WPI; 1999-190050/16.	
DR	N-PSDB; AAX28636.	
XX		
XX	New metalloprotease proteins ADAM16a and ADAM16b - useful in	
PT	contraception, and in the evaluation and treatment of fertility.	
PT		
XX		
PS	Claim 1; Fig 3A; 59pp; English.	
XX		
CC	This is a amino acid sequence of the novel membrane metalloprotease	
CC	protein designated ADAM 16a, isolated from human testis, and used in the	
CC	method of the invention. It may be used in contraceptive, vaccines, or in	
CC	fertility diagnosis and treatment. It may be involved sperm maturation,	
CC	or fertilisation through mechanisms such as proteolysis, cell adhesion,	
CC	cell fusion, or intracellular signalling	
XX		

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 27, 2004, 21:35:56 ; Search time 116 Seconds  
(without alignments)  
554.877 Million cell updates/sec

Title: US-10-633-202-2\_COPY\_399\_602

Perfect score: 1192

Sequence: 1 PFGNIFRLKYGKLVVEG.....QPHLNDTTCTGWDYHLGMAI 204

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1017041 seqs, 315518202 residues

Total number of hits satisfying chosen parameters: 1017041

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

SPTREMBL.25.\*  
1: sp\_archaea.\*  
2: sp\_bacteria.\*  
3: sp\_fungi.\*  
4: sp\_human.\*  
5: sp\_invertebrate.\*  
6: sp\_mammal.\*  
7: sp\_mhc.\*  
8: sp\_organelle.\*  
9: sp\_phage.\*  
10: sp\_plant.\*  
11: sp\_rodent.\*  
12: sp\_virus.\*  
13: sp\_vertebrate.\*  
14: sp\_unclassified.\*  
15: sp\_rvirus.\*  
16: sp\_bacteriap.\*  
17: sp\_archaeap.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	753	63.2	763	11 Q811Q4	Q811Q4 mus musculus
2	718.5	60.3	761	11 Q8CDV3	Q8CDV3 mus musculus
3	664	55.7	702	11 Q8BMR5	Q8BMR5 mus musculus
4	664	55.7	714	11 Q8K4K0	Q8K4K0 mus musculus
5	660	55.4	697	11 Q8BMR4	Q8BMR4 mus musculus
6	590.5	49.5	790	4 Q8TBZ7	Q8TBZ7 homo sapien
7	575	48.2	873	13 Q42595	Q42595 xenopus lae
8	572	48.0	845	11 Q81072	Q81072 mus musculus
9	572	48.0	853	11 Q8OU94	Q8OU94 mus musculus
10	567	47.6	655	4 Q8NFM6	Q8NFM6 homo sapien
11	519.5	43.6	732	11 Q811Q3	Q811Q3 mus musculus
12	517.5	43.4	703	11 Q8C0V3	Q8C0V3 mus musculus
13	507.5	42.6	754	11 Q8CGQ1	Q8CGQ1 mus musculus
14	502.5	42.2	473	11 Q80815	Q80815 mus musculus
15	502.5	42.2	763	11 Q8CGQ2	Q8CGQ2 mus musculus
16	501.5	42.1	735	11 Q60473	Q60473 cavia porce

17	489.5	41.1	751	11	P70535	P70535 rattus norv
18	485.5	40.7	474	6	Q951W7	Q951W7 macaca fasc
19	476	39.9	730	6	O19051	O19051 oryctolagus
20	472	39.6	706	13	O42593	O42593 xenopus lae
21	468.5	39.3	732	6	Q28484	Q28484 macaca fasc
22	467.5	39.2	713	6	Q28485	Q28485 macaca fasc
23	466.5	39.1	792	6	O19061	O19061 saguinus oe
24	466.5	39.1	919	6	Q28659	Q28659 oryctolagus
25	464	38.9	825	6	Q28477	Q28477 macaca fasc
26	464	38.9	838	6	O19056	O19056 papio anubi
27	464	38.9	905	6	Q28476	Q28476 macaca fasc
28	463.5	38.9	825	6	O46652	O46652 papio anubi
29	459.5	38.5	731	6	O19050	O19050 oryctolagus
30	451.5	37.9	821	6	O19060	O19060 saguinus oe
31	450	37.8	812	6	O77779	O77779 bos taurus
32	448	37.6	836	6	O19057	O19057 pongo pygma
33	444	37.2	804	11	Q60410	Q60410 cavia porce
34	440	36.9	203	11	P97777	P97777 rattus norv
35	438.5	36.8	899	13	Q8UVF1	Q8UVF1 coturnix co
36	428	35.9	155	6	O77634	O77634 sus scrofa
37	427	35.8	914	13	O12960	O12960 xenopus lae
38	426	35.7	735	6	Q866A8	Q866A8 sus scrofa
39	422.5	35.4	812	4	Q8N0W6	Q8N0W6 homo sapien
40	420	35.2	155	6	O77635	O77635 bos taurus
41	418	35.1	1023	5	Q967H9	Q967H9 strongyloce
42	413	34.6	922	13	Q8UVF2	Q8UVF2 coturnix co
43	407.5	34.2	694	5	Q9GZ15	Q9GZ15 drosophila
44	407.5	34.2	1182	5	Q9VKL1	Q9VKL1 drosophila
45	402.5	33.8	756	6	Q28483	Q28483 macaca fasc

## ALIGNMENTS

### RESULT 1

Q811Q4 ID Q811Q4 PRELIMINARY; PRT; 763 AA.  
AC Q811Q4;  
DT 01-JUN-2003 (Tremblrel. 24, Created)  
DT 01-JUN-2003 (Tremblrel. 24, Last sequence update)  
DT 01-OCT-2003 (Tremblrel. 25, Last annotation update)  
DE ADAM29.  
OS Mus musculus (Mouse).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
OX NCBI\_TaxID=10090;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=ICR; TISSUE=Testis;  
RA Choi I., Cho C.;  
RT "Cloning and characterization of ADAM29";  
RL Submitted (DEC-2002) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AY190759; AAC38663.1; -  
DR GO; GO:0004222; F:metalloendopeptidase activity; IEA.  
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.  
DR InterPro; IPR006586; ADAM cysteine.  
DR InterPro; IPR001762; Disintegrin.  
DR InterPro; IPR001590; Peptidase M12B.  
DR InterPro; IPR002870; Peptidase\_M12B\_N.  
DR Pfam; PF00200; disintegrin; 1.  
DR Pfam; PF01562; Pep\_M12B\_propep; 1.  
DR Pfam; PF01421; Reprolysin; 1.  
DR PRINTS; PR00289; DISINTEGRIN.  
DR ProDom; PD000564; Disintegrin; 1.  
DR SMART; SM00608; ACR; 1.  
DR SMART; SM00050; DISIN; 1.  
DR PROSITE; PS00215; ADAM\_MEPRO; 1.  
DR PROSITE; PS00427; DISINTEGRIN\_1; 1.  
DR PROSITE; PS00214; DISINTEGRIN\_2; 1.  
SQ SEQUENCE 763 AA; 86444 MW; F9EF082C3BFBA80 CRC64;  
Query Match 63.2%; Score 753; DB 11; Length 763;  
Best Local Similarity 58.6%; Pred. No. 5e-72;

```
Matches 119; Conservative 41; Mismatches 43; Indels 0; Gaps 0;
QY 2 YGNIFRLKYCNLVVEEGECDCGTIRQCAKDPCCLLNCTLHPGAACAFAFGICCKCKFL 61
Db 396 YTKDIFDRTRCGVGVVDEKQDCGSLRNCNTDLCMSNCTLTSGSSCAFGICCKNCQFL 455
QY 62 PSGLTLCRQGVGCDLPWCNCGTSHOCPPDDVYVQDGI SCNVNAFCYEKTCNNHDIQCKEIF 121
Db 456 PSGLTLCRKNDCIPLWCNCGTSHCPDDAYVEDGIPGVSAICYEKQCNDRNEHCRQIF 515
QY 122 GDARSASQSCYOEINTQGNRFGHCGIVGTTVVKWTPDIMCGRVQCNVGVIPNLIEHS 181
Db 516 GQNAKTASVHCYREINTKGRGCHGLOQPTTIYKCKSNDCALCGRIQCDNVVQIPNKKDHS 575
QY 182 TVQQPHLNDTTCWGTDYHLGMAI 204
Db 576 TIHFALVKNVSCWGTDTYHTGSL 598

RESULT 2
Q8CDV3 PRELIMINARY; PRT; 761 AA.
AC Q8CDV3;
DT 01-MAR-2003 (TRENBLrel. 23, Created)
DT 01-MAR-2003 (TRENBLrel. 23, Last sequence update)
DT 01-OCT-2003 (TRENBLrel. 25, Last annotation update)
DE A disintegrin and metalloprotease domain 24.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Testis;
RX MEDLINE=22354683; PubMed=12466851;
RA The FANTOM Consortium,
RT "Analysis of the mouse transcriptome based on functional annotation of
RT 60,770 full-length cDNAs.";
RL Nature 420:563-573 (2002).
DR EMBL; AK029528; BAC26497.1; -.
DR GO; GO:0004222; F:metalloendopeptidase activity; IEA.
DR GO; GO:0008270; F:zinc ion binding; IEA.
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR InterPro; IPR006586; ADAM_cysteine.
DR InterPro; IPR001762; Disintegrin.
DR InterPro; IPR006209; EGF like.
DR InterPro; IPR001590; Peptidase M12B_N.
DR InterPro; IPR002870; Peptidase M12B_N.
DR Pfam; PF00200; disintegrin_1.
DR Pfam; PF01562; Pep_M12B_propep; 1.
DR Pfam; PF01421; Reprolysin; 1.
DR ProDom; PD000664; Disintegrin; 1.
DR SMART; SM00608; ACR; 1.
DR SMART; SM00050; DISIN; 1.
DR PROSITE; PS50215; ADAM_MPRO; 1.
DR PROSITE; PS0214; DISINTEGRIN_2; 1.
DR PROSITE; PS01186; EGF 2; 1.
DR PROSITE; PS00142; ZINC_PROTEASE; 1.
SQ SEQUENCE 761 AA; 85039 MW; 2184A903262A3683 CRC64;

Query Match 60.3%; Score 718.5; DB 11; Length 761;
Best Local Similarity 55.6%; Pred. No. 2.5e-68;
Matches 114; Conservative 36; Mismatches 54; Indels 1; Gaps 1;
QY 1 PYPGNIFRLKYCNLVVEEGECDCGTIRQCAKDPCCLLNCTLHPGAACAFAFGICCKCK 59
Db 398 PPSPIFQLKVCNGVIVEEGECDCGSSCENRCNRCPPSCLTSKAKCDTGLCNRKQ 457
QY 60 FLPSGTLCRQGVGCDLPWCNCGTSHOCPPDDVYVQDGI SCNVNAFCYEKTCNNHDIQCKE 119
Db 458 IQPSGTLCRARENECDLPWCNCGTSHCPEDLVQDGTSCPGDGYCYEKRCKNSHDVHCOR 517
```

```
QY 120 IRGQDARSASQSCYOEINTQGNRFGHCGIVGTTVVKWTPDIMCGRVQCNVGVIPNLIE 179
Db 518 VFGQLAMKASQSCYKELNTRDRFGNCFINNEYVRCEISDILCGRIQCDKVGTLPILOQ 577
QY 180 HSTVQQPHLNDTTCWGTDYHLGMAI 204
Db 578 HYTIHWHFNSVSCWSTDYHLGMI 602

RESULT 3
Q8BMR5 PRELIMINARY; PRT; 702 AA.
AC Q8BMR5;
DT 01-MAR-2003 (TRENBLrel. 23, Created)
DT 01-MAR-2003 (TRENBLrel. 23, Last sequence update)
DT 01-OCT-2003 (TRENBLrel. 25, Last annotation update)
DE Similar to ADAM 26 precursor.
GN ADAM34.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Testis;
RX MEDLINE=22354683; PubMed=12466851;
RA The FANTOM Consortium,
RT "Analysis of the mouse transcriptome based on functional annotation of
RT 60,770 full-length cDNAs.";
RL Nature 420:563-573 (2002).
DR EMBL; AK029590; BAC26524.1; -.
DR MGD; MGI:2181992; Adam34.
DR GO; GO:0004222; F:metalloendopeptidase activity; IEA.
DR GO; GO:0008270; F:zinc ion binding; IEA.
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR InterPro; IPR006586; ADAM_cysteine.
DR InterPro; IPR001762; Disintegrin.
DR InterPro; IPR001590; Peptidase M12B_N.
DR InterPro; IPR002870; Peptidase M12B_N.
DR Pfam; PF00200; disintegrin; 1.
DR Pfam; PF01562; Pep_M12B_propep; 1.
DR Pfam; PF01421; Reprolysin; 1.
DR PRINTS; PR00289; DISINTEGRIN.
DR ProDom; PD000664; Disintegrin; 1.
DR SMART; SM00608; ACR; 1.
DR SMART; SM00050; DISIN; 1.
DR PROSITE; PS50215; ADAM_MPRO; 1.
DR PROSITE; PS00427; DISINTEGRIN_1; 1.
DR PROSITE; PS50214; DISINTEGRIN_2; 1.
DR PROSITE; PS00142; ZINC_PROTEASE; 1.
SQ SEQUENCE 702 AA; 79278 MW; 31976863CA4CE0EE CRC64;

Query Match 55.7%; Score 664; DB 11; Length 702;
Best Local Similarity 54.4%; Pred. No. 1.6e-62;
Matches 105; Conservative 33; Mismatches 55; Indels 0; Gaps 0;
QY 12 CGNLVVEEGECDCGTIRQCAKDPCCLLNCTLHPGAACAFAFGICCKCKFLPSGLTLCRQV 71
Db 398 CGNLVVEEGECDCGTESTCLQDPCSSDCVLKPGAQCAFGICCKNCQFLKTGTVCREK 457
QY 72 GECIDLPEWCNCGTSHOCPPDDVYVQDGI SCNVNAFCYEKTCNNHDIQCKEIFGDARSASQ 131
Db 458 NECIDLPEWCNCGTSABCPGDVYKADGIPCGSGEGCYKMECHQDEQCRKIFGNGSRADRI 517
QY 132 CYOEINTQGNRFGHCGIVGTTVVKWTPDIMCGRVQCNVGVIPNLIEHSTVQQPHLNDT 191
Db 518 CYMENNRGRDGRFGNCGNDSSMTTRCNLADVLGRIQCNVITQLPQRNHETVHYTHFSNV 577
QY 192 TCWGTDTYHLGMAI 204
```

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: August 27, 2004, 21:38:21 ; Search time 40 Seconds  
(without alignments)  
490.577 Million cell updates/sec

Title: US-10-633-202-2\_COPY\_399\_602  
Perfect score: 1152  
Sequence: 1 PYPGNIFRLKYGCVLVEEG.....QFHLNDTTCWGTVDYHLGMAI 204

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 283366 seqs, 96191526 residues

Total number of hits satisfying chosen parameters: 283366

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : PIR 78:\*  
1: pir1:\*  
2: pir2:\*  
3: pir3:\*  
4: pir4:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	567	47.6	655	JC7850	disintegrin and me
2	536.5	45.0	660	S71949	metalloproteinase
3	502.5	42.2	473	I43283	ADAM 4 protein pre
4	501.5	42.1	735	I48101	ADAM 6 protein pre
5	468.5	39.3	732	I52361	testicular metallo
6	467.5	39.2	713	I65253	disintegrin-like t
7	464	38.9	825	S55060	fertilin alpha-II
8	464	38.9	905	S55059	fertilin alpha-I -
9	442	37.1	600	I49281	fertilin alpha pre
10	425.5	35.7	903	S60257	metrin alpha - mo
11	417.5	35.0	151	S60259	metrin gamma - mo
12	402.5	33.8	756	S47656	tMDC II protein -
13	401.5	33.7	814	G02390	disintegrin-like m
14	391.5	32.8	777	I48100	ADAM 5 protein pre
15	390	32.7	734	JC4861	fertilin beta cha
16	386.5	32.4	776	S28258	androgen-regulated
17	384	32.2	670	I65967	disintegrin-like m
18	375	31.5	735	G02937	fertilin beta - cr
19	374	31.4	789	S28259	androgen-regulated
20	372	31.2	571	S24789	jararagin C preu
21	371	31.1	609	S52470	catrocollastatin p
22	369	31.0	419	A59414	metalloproteinase
23	367	30.8	150	S60258	metrin beta - mou
24	359	30.1	429	A42972	coagulation factor
25	357.5	30.0	357	S23403	sperm surface prot
26	356.5	29.9	952	T18900	disintegrin and me
27	354	29.7	617	S48160	metalloproteinase
28	353	29.6	610	JC7530	vascular apoptosis
29	347	29.1	524	S38539	disintegrin-like m

ALIGNMENTS

RESULT 1

JC7850  
disintegrin and metalloproteinase(ADAM) 9 protein, short form - human  
N:Alternate names: MDC9 protein; meltrin gamma  
C:Species: Homo sapiens (man)  
C:Date: 18-Nov-2002 #sequence\_revision 18-Nov-2002 #text\_change 31-Mar-2003  
C:Accession: JC7850  
R:Hotoda, N.; Koike, H.; Sasagawa, N.; Ishiura, S.  
Biochem. Biophys. Res. Commun. 293, 800-805, 2002  
A:Title: A secreted form of human ADAM9 has an alpha-secretase activity for APP.  
A:Reference number: JC7850; MUID:22050095; PMID:12054541  
A:Accession: JC7850  
A:Molecule type: mRNA  
A:Residues: 1-655 <HOT>  
A:Cross-references: GB:AF495383  
C:Comment: This protein, which is a member of the a disintegrin and metalloprotease (AD) lular functions. It is proteolytically active, and has an alpha-secretase activity for C:Genetics:  
A:Gene: adam9b

Query Match 47.6%; Score 567; DB 2; Length 655;  
Best Local Similarity 46.4%; Pred. No. 1.8e-36;  
Matches 97; Conservative 36; Mismatches 62; Indels 14; Gaps 4;  
QY 1 PYPGNIFRLKYGCVLVEEGECDCGTIRQCAKDPCL-LNCTLHPGACAFGICCKDCK 59  
Db 406 KPDEAYAPSCKGKLVDAEGECDCGTPECELDPCCEGSTCKLKSFAECAYGDCCKDCK 465  
QY 60 FLPSGTLRCQQVGECDLPWCNCGTSHQCPDDVYVQDGI SC-VNNAFCVYEKTCNNHDIQCK 118  
Db 466 FLPGGTLRCGKTSECDVPEYCNSSQFCQPDVFIQNGYPCQNNKAYCYNGMCQYDQACQ 525  
QY 119 EIFGDARASQSCVQIEINTQGNRFGHCGIVGTTTKVCTPDMCGRVQCEVY-----G 172  
Db 526 VIFGSKAAAPKDCFIENSKGDRFCNGCFSGNEKKCATGNALCGKLCENVQIEPVRG 585  
QY 173 VIPNLIEHSTVQQFHLNDTTCWGTVDYHLG 201  
Db 586 IVPAILIQTFS-----RGTKCWGVDFQLG 608

RESULT 2

S71949  
metalloproteinase 12 (EC 3.4.24.-) precursor - human  
N:Alternate names: disintegrin 12; myeloma cell metalloproteinase MCMP  
C:Species: Homo sapiens (man)  
C:Date: 19-Mar-1998 #sequence\_revision 17-Apr-1998 #text\_change 04-Feb-2000  
C:Accession: S71949; PC4264  
R:McKie, N.; Dallas, D.J.; Edwards, T.; Apperley, J.F.; Russell, R.G.G.; Croucher, P.I.  
Biochem. J. 318, 459-462, 1996  
A:Title: Cloning of a novel membrane-linked metalloproteinase from human myeloma cells.  
A:Reference number: S71949; MUID:96404892; PMID:8809033







GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: August 27, 2004, 21:35:21 ; Search time 24 Seconds  
(without alignments)  
442.596 Million cell updates/sec

Title: US-10-633-202-2\_COPY\_399\_602

Perfect score: 1192

Sequence: 1 PYPGNIFRLKYLGNLVVEG.....QFHLNDLTTCTGTDYHLGMAI 204

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 141681 seqs, 52070155 residues

Total number of hits satisfying chosen parameters: 141681

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt\_42.\*

Pred: No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	1192	100.0	726	1 AD20_HUMAN	O43506 homo sapien
2	864	72.5	722	1 AD21_HUMAN	Q9UKJ8 homo sapien
3	788	66.1	820	1 AD29_HUMAN	Q9UKF5 homo sapien
4	769	64.5	760	1 AD25_MOUSE	Q9R159 mus musculus
5	747	62.7	729	1 AD21_MOUSE	Q9J176 mus musculus
6	716.5	60.1	761	1 AD24_MOUSE	Q9R160 mus musculus
7	664	55.7	697	1 AD26_MOUSE	Q9R158 mus musculus
8	590.5	49.5	790	1 AD30_HUMAN	Q9UKF2 homo sapien
9	567	47.6	819	1 AD09_HUMAN	Q13443 homo sapien
10	461.5	38.7	789	1 AD01_RAT	P70505 rattus norv
11	450	37.8	791	1 AD1A_MOUSE	Q60813 mus musculus
12	442.5	37.1	909	1 AD12_HUMAN	O43184 homo sapien
13	437.5	36.7	956	1 AD19_HUMAN	Q9R013 homo sapien
14	425.5	35.7	903	1 AD12_MOUSE	Q61824 mus musculus
15	425.5	35.7	920	1 AD19_MOUSE	O35674 mus musculus
16	422.5	35.4	813	1 AD33_HUMAN	Q9B211 homo sapien
17	422	35.4	906	1 AD22_HUMAN	O9P0K1 homo sapien
18	421	35.3	935	1 AD22_XENLA	O42596 xenopus lae
19	420	35.2	806	1 AD1B_MOUSE	Q8R534 mus musculus
20	418	35.1	745	1 AD02_BOVIN	O77780 bos taurus
21	418	35.1	857	1 AD22_MOUSE	Q9R1V6 mus musculus
22	401.5	33.7	814	1 AD15_HUMAN	Q13444 homo sapien
23	399.5	33.5	754	1 AD07_HUMAN	Q9B2U9 homo sapien
24	399.5	33.5	864	1 AD15_MOUSE	O88839 mus musculus
25	395	33.1	751	1 AD02_RABIT	Q28660 oryctolagus
26	390	32.7	735	1 AD02_HUMAN	Q99965 homo sapien
27	387	32.5	773	1 AD11_MOUSE	Q9R1V4 mus musculus
28	386.5	32.4	776	1 AD07_MACFA	Q28475 macaca fasc
29	386.5	32.4	797	1 AD33_MOUSE	Q923W9 mus musculus
30	385	32.3	452	1 AD11_XENLA	Q9P233 xenopus lae
31	384	32.2	769	1 AD11_HUMAN	Q75078 homo sapien
32	380.5	31.9	816	1 AD15_RAT	Q9GVV0 r adam 15 p
33	377	31.6	788	1 AD07_MOUSE	O35227 mus musculus

RESULT 1

ID	AD20_HUMAN	STANDARD;	PRT;	726 AA.
AC	O43506; Q9UKJ9;			
DT	16-OCT-2001 (Rel. 40, Created)			
DT	16-OCT-2001 (Rel. 40, Last sequence update)			
DT	15-MAR-2004 (Rel. 43, Last annotation update)			
DE	ADAM 20 precursor (EC 3.4.24.-) (A disintegrin and metalloproteinase domain 20).			
GN	ADAM20.			
OS	Homo sapiens (Human).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.			
OX	NCBI_TaxID=9606;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	TISSUE=Testis;			
RX	MEDLINE=98137801; PubMed=9469942;			
RA	Hooft van Hujsduijn R.;			
RT	"ADAM 20 and 21; two novel human testis-specific membrane metalloproteinases with similarity to fertilin-alpha.";			
RL	Gene 206:273-282(1998).			
RN	[2]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=99453762; PubMed=10524237;			
RA	Poindexter K., Nelson N., DuBoise R.F., Black R.A., Cerretti D.P.;			
RT	"The identification of seven metalloproteinase-disintegrin (ADAM) genes from genomic libraries.";			
RL	Gene 237:61-70(1999).			
CC	- FUNCTION: May be involved in sperm maturation and/or fertilization.			
CC	- COFACTOR: Binds 1 zinc ion per subunit (Potential).			
CC	- SUBCELLULAR LOCATION: Type I membrane protein.			
CC	- TISSUE SPECIFICITY: Testis specific.			
CC	- DOMAIN: A tripeptide motif (VGE) within disintegrin-like domain could be involved in the binding to egg integrin receptor and thus could mediate sperm/egg binding.			
CC	- DOMAIN: The cysteine-rich domain encodes putative cell-fusion peptides, which could be involved in sperm-egg fusion.			
CC	- PTM: Has no obvious cleavage site for furin endopeptidase, suggesting that the proteolytic processing is regulated.			
CC	- MISCELLANEOUS: May be the functional equivalent of ADAM 1/fertilin alpha which is a pseudogene in human.			
CC	- SIMILARITY: Belongs to peptidase family M12B.			
CC	- SIMILARITY: Contains 1 EGF-like domain.			
CC	- SIMILARITY: Contains 1 disintegrin domain.			
CC	-----			
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <a href="http://www.isb-sib.ch/announce/">http://www.isb-sib.ch/announce/</a> or send an email to <a href="mailto:license@isb-sib.ch">license@isb-sib.ch</a> ).			
CC	-----			
CC	EMBL; AF029899; AAC52041.1; -.			

ALIGNMENTS

```

DR EMBL; AF158643; AAD55254.1; -.
DR HSSP; M12.218; 1FVL.
DR Genew; HGNC:199; ADAM20.
DR MIM; 603712; -.
DR GO; GO:0008237; F:metallopeptidase activity; TAS.
DR GO; GO:0007338; P:fertilization (sensu Animalia); TAS.
DR InterPro; IPR006586; ADAM cysteine.
DR InterPro; IPR001762; Disintegrin.
DR InterPro; IPR006209; EGF-like.
DR InterPro; IPR001818; Pept_M10A_M12B.
DR InterPro; IPR006025; Pept_M_Zn_BS.
DR InterPro; IPR001590; Peptidase_M12B.
DR InterPro; IPR002870; Peptidase_M12B_N.
DR Pfam; PF00200; disintegrin; 1.
DR Pfam; PF01562; Pep_M12B_propep; 1.
DR Pfam; PF01421; Reprolysin; 1.
DR PRINTS; PR00289; DISINTEGRIN.
DR ProDom; PD000664; Disintegrin; 1.
DR SMART; SM00608; ACR; 1.
DR SMART; SM00050; DISIN; 1.
DR PROSITE; PS50215; ADAM_MEPRO; 1.
DR PROSITE; PS00427; DISINTEGRIN_1; FALSE_NEG.
DR PROSITE; PS50214; DISINTEGRIN_2; 1.
DR PROSITE; PS00022; EGF_1; FALSE_NEG.
DR PROSITE; PS01186; EGF_2; FALSE_NEG.
DR PROSITE; PS00142; ZINC_PROTEASE; 1.
DR PROSITE; PS00546; CYSTEINE_SWITCH; FALSE_NEG.
KW Hydrolase; Metalloprotease; Zinc; Signal; Glycoprotein; Zymogen;
KW Transmembrane; EGF-like domain.
FT SIGNAL 1 31
FT PROPEP 32 206
FT CHAIN 207 726
FT DOMAIN 207 693
FT TRANSMEM 694 714
FT DOMAIN 715 726
FT DOMAIN 207 384
FT DOMAIN 407 493
FT DOMAIN 494 634
FT DOMAIN 635 663
FT SITE 173 173
FT SITE 469 471
FT ACT_SITE 343 343
FT METAL 342 342
FT METAL 346 346
FT METAL 352 352
FT DISULFID 357 379
FT DISULFID 465 478
FT DISULFID 635 646
FT DISULFID 640 652
FT DISULFID 654 663
FT CARBOHYD 191 191
FT CARBOHYD 226 226
FT CARBOHYD 378 378
FT CARBOHYD 438 438
FT CARBOHYD 479 479
FT CARBOHYD 587 587
FT CONFLICT 19 19
FT CONFLICT 109 109
FT CONFLICT 637 637
FT CONFLICT 726 726
FT CONFLICT 817 817
SQ SEQUENCE 726 AA; 81710 MW; BF356A6771CD1C12 CRC64;

Query Match 100.0%; Score 1192; DB 1; Length 726;
Best Local Similarity 100.0%; Pred. No. 2.8e-90;
Matches 204; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PYPGNIFRLKYCGNLVWEGEGCDGTTIRQCAKDPCCLLNCTLHPGAACAFGICCKCKF 60
DB 399 PYPGNIFRLKYCGNLVWEGEGCDGTTIRQCAKDPCCLLNCTLHPGAACAFGICCKCKF 458
QY 61 LPSGTLCRQQVQCECDLPEWNCNGTSHQCDDVYVQDGIISCNVNAFCYKTCNNHDIQCKEI 120

```

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: June 15, 2005, 20:15:51 ; Search time 1341 Seconds  
(without alignments)  
3204.871 Million cell updates/sec

Title: US-10-633-202-2

Perfect score: 4030

Sequence: 1 MAVGRLVHVRVTLILLWFG.....LFLHLVLPKRTKXEDDEG 726

Scoring table: BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 4390206 seqs, 2959870667 residues

Total number of hits satisfying chosen parameters: 8780412

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DEV=xlp  
-O=/cpn2.1/USPTO.spool.p/US10633202/runat.14062005.141725.13079/app.query.fasta\_1.903  
-DB=N\_Geneseq.16Dec04 -QFMT=fastap -SUFFIX=ring -MINMATCH=0.1 -LOOPEL=0  
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MAIRIX=biosum62 -TRANS=human40.cdi  
-LIST=45 -DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=25  
-MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000  
-USER=US10633202 @CGN.1.1.708 @runat.14062005.141725.13079 -NCPU=6 -ICPU=3  
-NO\_WMAP -LARGEQUERY -NEG\_SCORES=0 -WAIT -DSPLOCK=100 -LONGLOG  
-DEV\_TIMEOUT=120 -WARN\_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6  
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : N\_Geneseq.16Dec04:\*  
1: Geneseqn1980s:\*  
2: Geneseqn1990s:\*  
3: Geneseqn2000s:\*  
4: Geneseqn2001as:\*  
5: Geneseqn2001bs:\*  
6: Geneseqn2002as:\*  
7: Geneseqn2002bs:\*  
8: Geneseqn2003as:\*  
9: Geneseqn2003bs:\*  
10: Geneseqn2003cs:\*  
11: Geneseqn2003ds:\*  
12: Geneseqn2004as:\*  
13: Geneseqn2004bs:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	4030	100.0	2181	2 AAX56461	Aax56461 Human SVP
2	4030	100.0	3230	8 ABZ76248	Abz76248 Human GEN
3	4004	99.4	2426	2 AAX28636	Aax28636 Nucleotid
4	4004	99.4	2442	10 ADC78876	Adc78876 Human PRO
5	4004	99.4	2442	10 ADR85180	Adr85180 Farnesyl

6	2435	60.4	2169	2 AAX90699	Aax90699 Human SVP
7	2435	60.4	2169	8 ABX11650	Abx11650 Human cDN
8	2102.5	52.2	2301	3 AAX59304	Aax59304 DNA encod
9	2102.5	52.2	2364	3 AAX59305	Aax59305 DNA encod
10	2102.5	52.2	2463	3 AAX59306	Aax59306 DNA encod
11	2102.5	52.2	2629	10 ADC19803	Adc19803 Human can
12	2102.5	52.2	4007	10 ADC19811	Adc19811 Human can
13	2102.5	52.2	5112	10 ADC19808	Adc19808 Human can
14	2102.5	52.2	5165	10 ADC19793	Adc19793 Human can
15	2102.5	52.2	6106	10 ADC19805	Adc19805 Human can
16	2102.5	52.2	6159	10 ADC19792	Adc19792 Human can
17	2102.5	52.2	6160	10 ADC19802	Adc19802 Human can
18	2102.5	52.2	6160	10 ADC19798	Adc19798 Human can
19	2102.5	52.2	6213	10 ADC19812	Adc19812 Human can
20	2102.5	52.2	6213	10 ADC19809	Adc19809 Human can
21	2102.5	52.2	6214	10 ADC19795	Adc19795 Human can
22	2102.5	52.2	6267	10 ADC19806	Adc19806 Human can
23	1717.5	42.6	1856	2 AAX28637	Aax28637 Nucleotid
24	1654	41.0	1653	4 AAD21436	Aad21436 ADAM-20d1
25	1387.5	34.4	2373	3 AAX59226	Aax59226 CDNA enco
26	1387.5	34.4	2694	10 ADC78910	Adc78910 Human PRO
27	1387.5	34.4	2695	4 AAX21345	Aax21345 Human cDN
28	1387.5	34.4	2695	8 ACA03704	Aca03704 CDNA enco
29	1387.5	34.4	2695	8 ABX89242	Abx89242 DNA encod
30	1387.5	34.4	2695	8 ACD41896	Acd41896 Human sec
31	1387.5	34.4	2695	8 ACA04125	Aca04125 Human cDN
32	1387.5	34.4	2695	9 ADA45722	Ada45722 Novel hum
33	1387.5	34.4	2695	9 ADA76153	Ada76153 Human PRO
34	1387.5	34.4	2695	9 ADA18803	Ada18803 Human PRO
35	1387.5	34.4	2695	9 ADA61426	Ada61426 Homo sapi
36	1387.5	34.4	2695	9 ADB19211	Adb19211 Novel hum
37	1387.5	34.4	2695	9 ADB27752	Adb27752 CDNA enco
38	1387.5	34.4	2695	9 ADA86231	Ada86231 Novel hum
39	1387.5	34.4	2695	9 ADB15795	Adb15795 Human PRO
40	1387.5	34.4	2695	9 ADA47581	Ada47581 Human PRO
41	1387.5	34.4	2695	9 ADA67376	Ada67376 Human PRO
42	1387.5	34.4	2695	9 ADB30383	Adb30383 CDNA enco
43	1387.5	34.4	2695	9 ADA85679	Ada85679 Novel hum
44	1387.5	34.4	2695	9 ADA96891	Ada96891 Human PRO
45	1387.5	34.4	2695	9 ADA79195	Ada79195 Human PRO

#### ALIGNMENTS

RESULT 1  
AAX56461  
ID AAX56461 standard; DNA; 2181 BP.  
XX  
AC AAX56461;  
XX  
DT 26-JUL-1999 (first entry)  
XX  
DE Human SVPHI-26 protein encoding DNA.  
XX  
KW Human; SVPHI-26; proteinase; testis; fertilisation; spermatogenesis;  
XX birth control; detergent additive; diagnosis; testicular cancer; ss.  
XX Homo sapiens.  
XX  
XX WO9923228-A1.  
XX  
XX PD 14-MAY-1999.  
XX  
XX PF 30-OCT-1998; 98WO-US022965.  
XX  
XX PR 30-OCT-1997; 97US-0063571P.  
XX  
XX PA (IMMUNEX CORP.  
XX  
XX PI Cerretti DP;  
XX  
XX DR WPI; 1999-337487/28.  
XX  
XX DR P-PSDB; AAY17413.

XX New human SVPH1-26 DNA useful for the diagnosis and prognosis of  
PT testicular cancers.  
PT

PS Claim 1; Page 70-73; 96pp; English.

The present sequence encodes human SVPH1-26. Human SVPH1-26 polypeptides are proteinases implicated in fertilisation and spermatogenesis. They can be used as therapeutic agents. A proteinase inhibitor of the catalytic domain would inhibit SVPH1-26 activity and would be useful as a method for birth control. Also, an inhibitor of the disintegrin domain of SVPH1-26 may affect fertilisation. The proteinase activity of SVPH1-26 can also be used as a detergent additive for the removal of strains having a protein component. The SVPH1-26 polypeptides and fragments can also be used as molecular weight markers, as markers for determination of isoelectric points of sample proteins and as controls for establishing the extent of fragmentation of a protein sample. The products can also be used for identifying, separating or purifying cells that express SVPH1-26 polypeptides such as testis cells. They can be used for the diagnosis and prognosis of testicular cancers.

Sequence 2181 BP; 600 A; 423 C; 530 G; 628 T; 0 U; 0 Other;

**Alignment Scores:**

[illegible]

US-10-633-202-2 (1-726) x AAX56461 (1-2181)

[illegible]